Project Report: Customer Lifetime Value (LTV) Prediction Model

Prepared by: Nikitha V

1. Introduction

Customer Lifetime Value (LTV) is a critical metric for businesses to predict the long-term revenue potential of customers. This project focuses on building a machine learning model to predict LTV using historical transaction data. The goal is to help businesses identify high-value customers, optimize marketing spend, and improve retention strategies.

2. Abstract

This project involves:

- Data preprocessing (cleaning, feature engineering).
- RFM analysis (Recency, Frequency, Monetary Value).
- Training an XGBoost model to predict LTV.
- Segmenting customers into High/Medium/Low LTV tiers.

3. Tools Used

- Python (Pandas) - Data cleaning & manipulation
- Python (Scikit-learn) - Machine learning (train-test split, metrics)
- XGBoost - Predictive modeling
- Matplotlib/Seaborn - Data visualization
- Jupyter Notebook - Code execution & documentation.

4. Steps Involved

Step 1: Data Collection & Cleaning

- Used a retail transactions dataset (e.g., from Kaggle).

- Removed duplicates, handled missing values, and calculated `total_spend` per transaction.

Step 2: Feature Engineering (RFM Analysis)

- Recency (R): Days since the last purchase.
- Frequency (F): Total number of orders.
- Monetary Value (M): Total revenue per customer.
- Average Order Value (AOV): `Total Spend / Frequency`.

Step 3: Model Training

- Split data into 80% training and 20% testing.
- Trained an XGBoost regressor to predict LTV.
- Evaluated performance using MAE and RMSE.

Step 4: Visualization & Segmentation

- Created feature importance plots (most impactful: [Top Feature]).
- Segmented customers into High/Medium/Low LTV tiers using quantiles.

Step 5: Exporting Results

- Saved predictions as a CSV file (`customer_ltv_predictions.csv`).

5. Conclusion

This project successfully:

- Predicted LTV with reasonable accuracy
- Highlighted key drivers of customer value (e.g., Frequency, AOV).
- Enabled actionable segmentation for targeted marketing.